Agricultural Plant Leaf Disease Prediction Using Deep Learning Models

Synopsis of the Research Proposal

Submitted for the Registration in the

Ph.D. Program (Part-Time) in Engineering

By

MEGHA RATHORE

Under Supervision of

DR. MUKESH KALLA

Associate Professor, ICFAI University

THE ICFAI UNIVERSITY JAIPUR

Proposed Summary / Abstract

Plants are indispensable not only for food production but also for their applications in medicine and the environment. However, plant diseases, caused by various pathogens like fungi, bacteria, and viruses, significantly reduce crop yields and quality, posing a threat to food security in a country where millions still face hunger.

The proposed research aims to develop an automated plant disease detection system using machine learning (ML) and deep learning (DL) techniques. Current methods of disease detection rely heavily on manual inspection, which is time-consuming, error-prone, and inaccessible to many rural farmers. Leveraging Convolutional Neural Networks (CNNs) and transfer learning techniques, this study seeks to create a system capable of accurately classifying and detecting plant diseases from leaf images.

The research begins with the acquisition and preprocessing of image data to enhance quality and reduce noise, followed by segmentation to isolate diseased areas. Feature extraction techniques will identify critical patterns, such as discolorations or structural changes in leaves. The system will then employ hybrid CNN models optimized with transfer learning to enhance accuracy, even with limited datasets. Data augmentation will further address the scarcity of annotated data by generating synthetic samples, improving the model's robustness across diverse environmental conditions.

The proposed system offers several advantages, including scalability, high accuracy, and the potential for real-time application via smartphone integration or drone-based monitoring. By automating disease detection, the system will reduce dependency on expert diagnosis, lower operational costs, and provide timely interventions, ultimately increasing crop yield and quality. The research also highlights the significance of early detection in mitigating economic losses, which annually amount to billions in India due to plant diseases.

This study aligns with sustainable agricultural practices and has implications for global food security. Its findings will contribute to precision farming and the efficient management of plant health, leveraging technology to address the challenges of biodiversity loss, disease outbreaks, and environmental degradation. The research underscores the transformative potential of integrating artificial intelligence into agriculture, paving the way for innovative solutions to age-old challenges.

1. INTRODUCTION

1.1 Plant Species

One of the most important areas of research is the study of plant diseases and how to identify them quickly. Since India is a developing country, most Indians are employed in agriculture. Food security is ensured because the majority of people rely on agriculture for their livelihood. In addition to being necessary for human existence, plants also serve a number of useful purposes in the medical profession and other sectors. Humans and every other living things on Earth depend on plants to survive. Therefore, knowing how to automatically identify plant species and their medicinal applications is crucial.

Agriculture accounts up a large portion of India's domestic output, indicating the sector's importance to the national economy. India is one of the top three nations in the world for producing a wide range of agricultural products, including rice, wheat, cotton, and other vegetables. Plants may also suffer from a variety of diseases. The rapid and automated identification of illnesses is essential to preventing damage to the crops. A machine learning technique—more especially, deep learning techniques—will be helpful for automatic identification.

This thesis introduces automated plant disease detection. We were able to identify the species from the leaves since they stay on the plant throughout the year. On the other hand, fruits and blossoms only appear at specific seasons of the year. Since the disease's symptoms first appear on the leaves of plants, we have also considered these while diagnosing plant diseases. A more thorough description of plants and plant diseases follows this introduction [1].

- The plant has an important role in the Earth's atmosphere in addition to the industrial and medicinal domains. The plant is necessary to balance nature and maintain a healthy, breathing atmosphere.

- Humans and the natural world are closely related as they depend on environmental components for necessities like food and medication. Medicinal plants play important functions in both traditional medicine and as a commodity in international trade [2]. For the record, more than three quarters of the world's population rely on plants as their main source of medication. According to a recent World Health Organization (WHO)

survey, the vast majority of individuals use herbal treatments medication to enhance their well-being, and according to WHO estimates, some 21,000 plant species have medicinal use. Many medicinal plants that are frequently utilized in traditional medicine may be found in India, including

Neem's purifying properties have been used by Ayurvedic practitioners for thousands of years. Its effects on the circulatory, pulmonary, and urinary systems, as well as during pregnancy, have proven very beneficial.

Indian Long Pepper is a powerful stimulant of the respiratory and digestive systems that also revitalizes the lungs. It's significant when it comes to the release of metabolic heat energy. This effect is brought on by elevated blood levels of thyroid hormone.

• One important spice that improves the efficiency of the digestive system and the body's ability to absorb nutrients from meals is black pepper.

• Ginger, a common spice, is used in Ayurvedic medicine to promote digestion and prevent illness. Bowel movements are made easier and the muscles controlling digestion are relaxed as a result of these qualities.

• The body uses mint leaves' anti-oxidant, anti-inflammatory, and anti-diarrheal qualities to combat oxidative stress, gastrointestinal disorders, and the common cold.

• Rich in vitamin C, lemon leaves enhance circulation and give you a quick energy boost. There are several cosmetic applications for its oil extract. Additionally, it may help reduce insomnia and other stress-related symptoms..

1.2 Leaf Geometric Structure

The flattened, often green leaf is the main organ in charge of photosynthesis, according to plant taxonomy [3]. All leaves may be divided into two categories, as shown in Figure 1.1: simple leaves, which have just one leaflet, and complex leaves, which have several leaflets. For species classification, a plant's leaf border is essential. The leaf structure differs, as seen in Figure 1.2. Botanists and agronomists can classify plant species into several groups based on the properties of their leaves, such as size, shape, vein pattern, number of lobes, kind of edge, etc. Plant identification is important for a number of reasons, such as the fact that plants are the most basic food source for all living creatures, that they may have

therapeutic uses, and that the distribution of plants is changing due to our changing environment. Conventional plant identification techniques are time-consuming, intricate, and sometimes perplexing to non-experts.

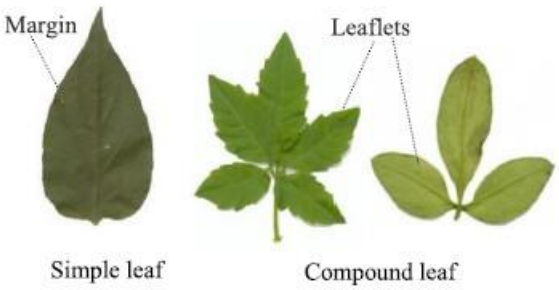


Figure 1.1: Leaf categories [6]

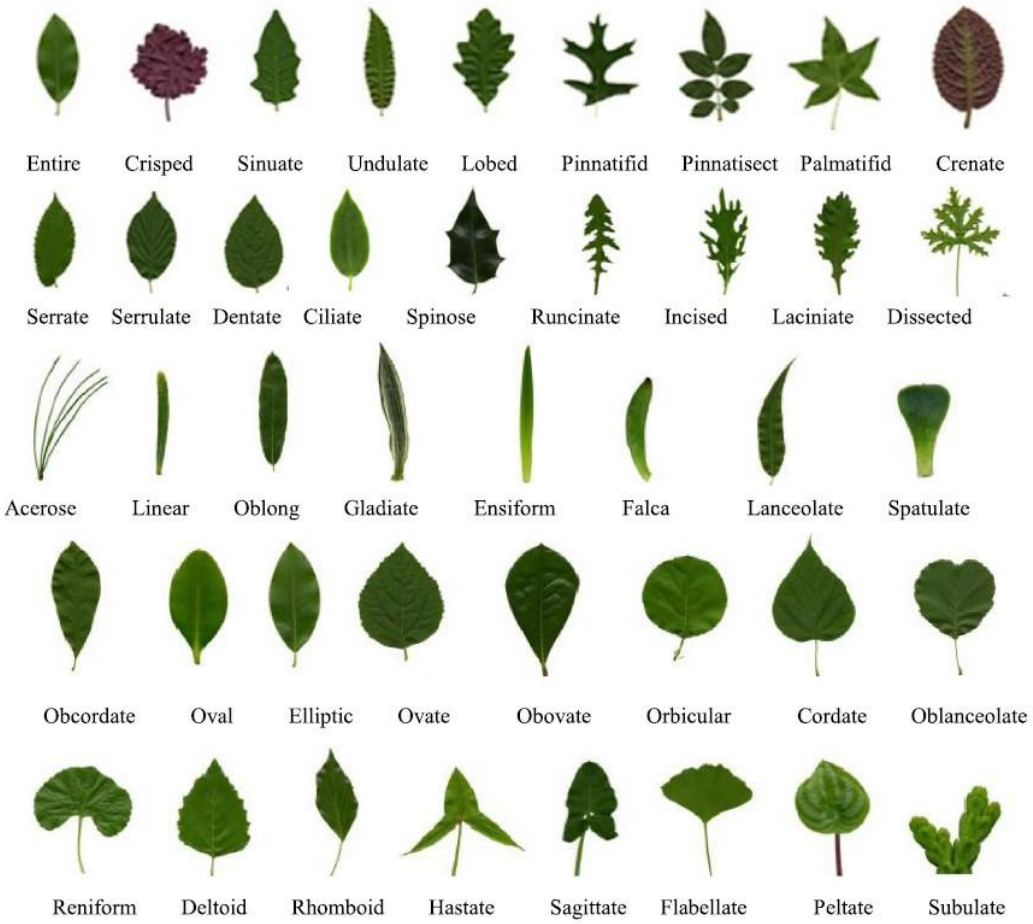


Figure 1.2: Structure of leaves [6]

1.3 Need for Automated Identification of Plants

In recent years, automated species identification using plant leaves has become a crucial component due to the rapid reduction of biodiversity and the skill of people working in plant taxonomy. Another concern is that, as a result of environmental deterioration, many rare plant species are already extinct, but many more are in danger of doing so. Thus, studies on plant identification can support greenwashing initiatives. The plant and animal worlds are constantly changing due to human activity and other factors, and the extinction of many species is a real possibility. It also recognizes the significance of understanding natural systems that sustain biodiversity.

1.4 Diseases in Plants

A plant disease is defined as something that "impedes a plant from functioning to its maximum potential". Plants are considered susceptible to infection when external factors disrupt their physiological processes, altering their growth, structure, functions, or other traits. Diseases and pests damage crops and interfere with the physiological functions of plants. A variety of plant diseases, including potato virus, leaf blight, leaf blast, and early blight, are depicted in Figure 1.3.

1.4.1 Factors Affecting Plant Diseases

Environmental factors have the power to either promote healthy plant development or inhibit it. Each plant disease and its effects on plants have a preferred temperature range. Consider the tobacco mosaic virus. It generates distinct, little, necrotic lesions on tobacco plants when it infects them at low temperatures. In winter, when temperatures are low, potato blight diseases are common.

- A leaf's relative humidity affects the development of rots and fungal diseases. When humidity levels are too high, bacteria, fungi, and water molds destroy fruits and foliage. The Rhizopus soft rot disease of sweet potatoes is one type of rot disease that might arise from high humidity levels. The growth of tomato leaf mold and powdery mildew spores may occur when the relative humidity is between 90% and 95%.

- Too much or too little soil moisture might affect the diseases caused by root rot. Dry soil causes fungal illnesses such as onion white rot, common scab in potatoes, and charcoal rot

in maize and soybeans.

• Fertility of the Soil: Diseases such as northern leaf blight of maize, powdery mildew of wheat, fire blight of apples and pears, and stalk rots of maize were caused by infectious diseases that reduced nutrient levels.

1.4.2 Various Diseases in Plants

Numerous pathogens, such as bacteria, fungus, and viruses, can cause plant diseases. A symptom is an obvious indication of a plant disease. The plant may show signs including a visible change in color, shape, or function as a reaction to the infection. Some of the most common plant diseases and their symptoms include the following:

Fungal illnesses: One of the most frequent causes of plant illnesses is fungus. They may infect a broad range of plant species and flourish in warm, humid conditions.

• Powdery mildew: A kind of fungus that causes a white, powdery growth on plant leaves, stems, and buds is the source of this illness. It impacts a variety of plants, such as grapes, cucumbers, and roses. Stunted growth can result from powdery mildew's restriction of photosynthesis and potential to damage the plant.

• Rusts: Plants including wheat, barley, and other crops are impacted by rust fungus. If left unchecked, these fungi's rust-colored pustules on the leaves can cause defoliation, decreased production, and possibly plant death.

• Late Blight: A serious condition affecting tomatoes and potatoes, late blight is brought on by the fungus Phytophthorainfestans. In chilly, damp weather, it causes black, water-soaked lesions on leaves and stems that can spread quickly and cause crop losses.

Bacterial Diseases: Plants may become infected with bacteria, which frequently results in withering, discoloration, and decay. Usually, insects or water are the means by which these illnesses are transmitted.

• Bacterial Wilt: This disease, which is brought on by Ralstoniasolanacearum, affects a variety of plants, such as potatoes, tomatoes, and peppers. The illness causes the leaves to wilt and turn yellow, which ultimately causes the plant to collapse.

• Canker: Fruit trees and other woody plants are susceptible to this bacterial disease. It results in cankers or sunken lesions on the bark, which can girdle branches and produce dieback.

Viruses are tiny diseases that may cause a range of symptoms in plants, including discolouration, curled leaves, and restricted development.

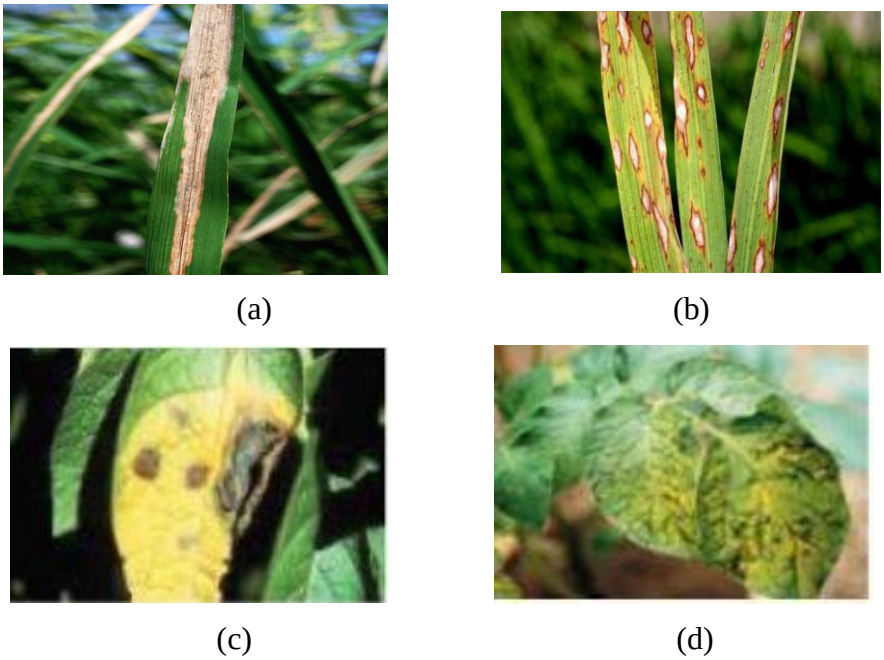
• Tomato Mosaic Virus (TMV): Among the most well-known plant viruses, TMV reduces

photosynthetic efficiency and results in poor fruit production by causing mosaic-like patterns of green and yellow on tomato leaves.

• Cucumber Mosaic Virus (CMV): This virus causes leaf curl, yellowing, and reduced development in cucumbers, peppers, and other crops. It can significantly reduce agricultural output and is spread by aphids.

Nematode diseases are caused by tiny worms called nematodes that infest plant roots. By feeding on root tissues, they can seriously harm plants, resulting in inadequate absorption of nutrients and water.

• Root Knot Nematodes: These nematodes obstruct the absorption of nutrients and water by forming galls or "knots" on plant roots. This causes growth retardation, wilting, and yellowing, particularly in crops like potatoes, cotton, and tomatoes.



(a) (b)

(c) (d)

Figure 1.3: The sample image of the leaf affected by (a) Leaf blight (b) Leaf blast (c) Early blight (d) Potato virus [6]

1.5 Multiple Leaf Disease

In most nations, agriculture is unquestionably the primary source of revenue and a significant component to GDP. Similar to people, plants may have a variety of diseases. Plant diseases have an adverse effect on crop quantity and quality. Plant diseases have a major impact on crop output. According to a research by the Associated Chambers of

Commerce and Industry of India, the annual agricultural losses resulting from pests and illnesses amount to Rs. 50,000 crore (500 billion), which is noteworthy given that at least 200 million Indians go to bed hungry every night. Therefore, the study of plants is quite useful. Timely detection of diseases is crucial to ensuring crop quality and product production.

It might be challenging for even a skilled agriculturist or pathologist to identify a plant disease based on the wide variety of crops that are farmed. However, in rural areas of developing countries, visual inspection continues to be the primary means of identifying illnesses. It also has to be regularly monitored by experts. For farmers in remote areas, traveling long distances to consult an expert might be expensive and time-consuming. Agronomists and farmers benefit from the high throughput and precision of automated computing systems for the diagnosis and detection of plant diseases. The impacted leaf samples are displayed in Figure 1.4.

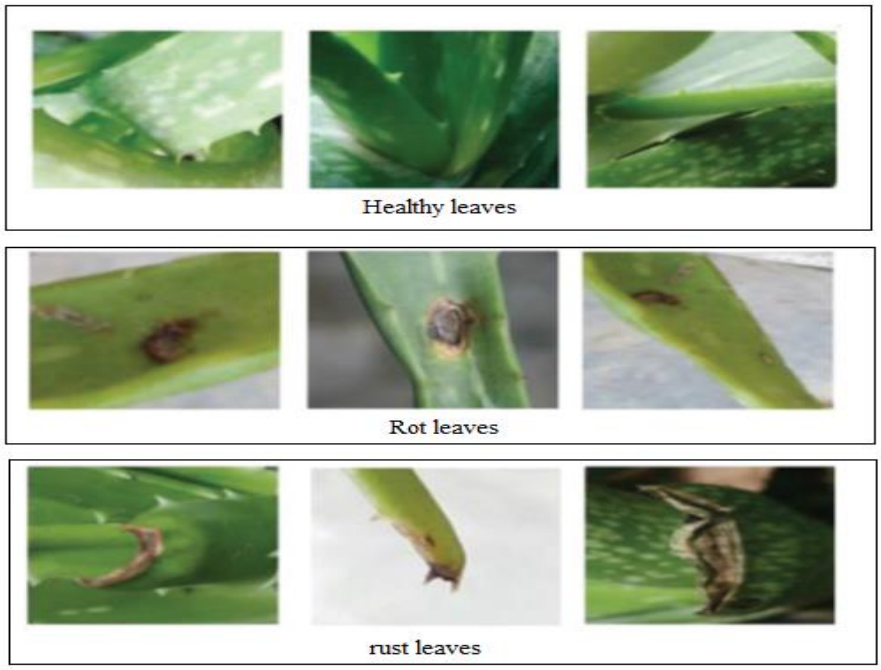


Figure 1.4: Samples of leaves [1]

In this instance, the top row displays healthy leaves, followed by rotten leaves in the second row and rusty leaves in the third. First of all, leaf plants are resilient. Even while the aloe plant is hardy, it does occasionally get ill.

To begin with, leaf plants are quite resilient. The aloe plant is susceptible to sporadic

illnesses despite its resilience.

Figure 1.5 shows how "aloe rust" affects succulent aloe plants. Circular, dark brown, or even black patches on leaves can be caused by fungus. The localized fungal infection goes away on its own.

Figure 1.5: Samples of Aloe Rust [24]

A fungal illness known as sooty mold appears in figure 1.6 following an aphid or mealy bug infestation. When pests like mealybugs and aphids deprive plants of water, they leave behind sticky, transparent liquids called honeydew on the leaves. The humid circumstances brought on by the honeydew encourage the growth of sooty mold.

Figure 1.6: Samples of Sooty Mold [24]

Figure 1.7 illustrates how basal stem rot may result from either too much moisture or too low of a temperature. Stems degrade as a result of this problem. The plant's underside usually sustains the most damage. When basal stem rot is present, rotted aloe tissues become black or reddish brown. By cutting a stem just above the rotten area, you can save some of the plant.

Figure 1.7: Samples of Basal Stem Rot [24]

However, plant diseases have a detrimental effect on productivity, food quality, and economic output. The many leafscale is the most well-known plant in contemporary industry and medicine [1]. Originally from central-southern Africa, arborescent aloe Miller, often called "candelabra Aloe," is a popular decorative and therapeutic plant in South Africa, Asia, Russia, and Japan [4]. Inflammation, cancer treatment, diabetes, hepatitis, inflammatory bowel disease, high cholesterol, asthma, osteoarthritis, stomach ulcers, insect bites, fever, general tonic, skin issues, hair issues, and radiation-induced skin sores are a few of the ailments for which this use is thought to be more important. People have historically used various leaf remedies to treat wounds, burns, and inflammation. Additionally, among other medicinal properties, certain leaves include antidiabetic, antioxidant, antihyperlipidemic, and anticancer properties. Vitamins A, C, E, and B12; enzymes like amylase, catalase, and peroxidase; minerals like zinc, copper, selenium, and calcium; sugars like mannose-6-phosphate and polysaccharides like glucomannans; anthraquinones like aloin and emodin; fatty acids like salicylic acid; lignin; and saponins are some of the more than 75 substances that have been identified in multiple leaves. [8]. These kinds of fungi usually infect fully developed plants, and the first places symptoms appear are on the undersides of the leaves. According to the diagnosis, anthracnose is a disease that targets and kills both early brilliant and normal seeds. Soft decay is caused by a fungal infection. The most prevalent type of fungus that may rapidly destroy a crop while it is only beginning to bloom and grow vegetatively is macrophominaphaseolina.

Aloe (Aloe vera), a lily plant, has long been used for its therapeutic qualities. Its natural habitats are in Africa and the Mediterranean region. There are rumors that it grows wild in the islands of Malta, Sicily, Cyprus, Carary Cape, and Cape Verde, as well as in the deserts

of India. It is a tropical perennial that can grow even in regions with limited rainfall. However, its full potential has not yet been realized. Although aloe is considered "a new botanical resource with the most promising potential in the world," it is all too often overlooked. Only a few isolated locations along the southern Indian coast are home to its native habitats. The major producers and exporters of aloe products include the United States, China, Mexico, Australia, and a few South American nations. Many nations employ the plant because of the expanding nutraceutical and cosmetics industries. Aloe is increasingly being used in the cosmetics sector as a safe and natural alternative to artificial chemicals. The Indian economy as a whole is significantly influenced by the greatest economic area in terms of population. With 159.7 million hectares, India has the second-largest arable land area in the world, behind the United States. With 82.6 million hectares, it has the largest gross irrigated agricultural area of any nation in the world [11]. India accounts for between 50 and 60 percent of global agricultural production. Generally speaking, one hectare of land should provide 15–20 tons of aloe vera. The output per acre might increase to 30 to 35 with proper irrigation and management. On the other hand, researchers estimate that the value of the aloe vera product market in India will reach $242 million by 2022. [13]. The cost analysis of aloe vera production in India is displayed in Table 1.1.

Table 1.1 Cost Analysis of Production of Aloe Vera in INR [62]

Sr. No.	Item	Year 1	Year 2	Year 3 onwards
I	**MATERIALS**			
1	Planting materials	28,000		
2	FYM	5,000	5,000	5,000
3	Fertilizers	1,900	1,900	1,900
4	Irrigation	1,250	1,250	1,250
5	Plant protection chemicals	1,250	1,250	1,250
	Sub Total – I	37,650	9,400	9,400
II	**LABOUR**			
1	Land preparation	4,800		
2	Preparation of ridges and furrows and planting	4,800		
3	Application of manures and fertilizers	650	650	650

4	Spraying of plant protection chemicals	650	650	650
5	Irrigation	950	950	950
6	Inter culture	3,200	3,200	3,200
7	Harvesting	-	3,200	4,800
8	Packing, loading etc.		1,700	3,200
	Sub Total – II	15,050	10,350	13,450
	GRAND TOTAL	52,700	19,750	22,850

Figure 1.8: Healthy Leaf Images [24]

The plant requires its leaves to survive. The degree of disease-induced leaf damage is one of the most important aspects of plant growth. Illnesses affect a variety of plant functions, such as transpiration, photosynthesis, germination, pollination, and many more. Therefore, early detection of leaf (foliar) diseases is necessary to increase agricultural production [1].

The healthy leaf (figure 1.8), which is distinguished by its immaculate green hue and lack of disease, is the first of two types of leaves. Bacteria, viruses, and fungi are among the pathogenic organisms that are wary of the second kind of leaf. The disease changes the original color, texture, and size of the leaf. The leaf has a lush hue. The new leaf is as thin as a violin, a sharp contrast to the old, extremely dark green one.

A healthy leaf is a key component of photosynthesis, which is a direct sign of plant development. Photosynthesis is essential to the existence and persistence of plants. A verdant leaf is a sign of a healthy plant. The leaves of the plant are very important. Numerous things, including as inadequate nutrition levels, illnesses, pests, shadow,

flooding, and more, can harm or damage leaves. Therefore, reducing leaf (foliar) illnesses as soon as feasible is thought to be an essential step in raising crop productivity. Growth, infections, and microorganisms are some of the most important factors that cause illness. There is actual agricultural damage brought on by these diseases. In terms of the bacteria they impact, these organisms have shown to be the most widespread infections.

1.6 Motivation

Deep learning techniques play a significant role in agriculture**,** particularly in plant leaf disease prediction. By leveraging computer vision, deep neural networks, and AI-driven models, farmers and agricultural experts can detect and diagnose plant diseases early, leading to better crop management and higher yields. Here's how deep learning is useful in this field:

Automated Disease Detection: Deep learning models, especially Convolutional Neural Networks (CNNs)**,** can analyze images of plant leaves to detect diseases, pests, and nutrient deficiencies with high accuracy.

- **Example:** CNNs trained on datasets of healthy and diseased plant leaves can classify and identify diseases such as powdery mildew, rust, or bacterial blight.

High Accuracy in Disease Identification: Traditional methods rely on manual inspection, which can be time-consuming and error-prone. Deep learning models can achieve higher accuracy rates (>90%) by learning complex patterns in leaf textures, colors, and shapes.

- **Example Models:**
    - ResNet, VGG, Inception **–** Used for feature extraction and classification.
    - YOLO (You Only Look Once) – Useful for real-time detection in the field.

Early Disease Detection and Prevention: Deep learning helps in early-stage disease detection**,** allowing farmers to take corrective actions before the disease spreads**.**

- **Example:** Using hyperspectral imaging and deep learning, minor discolorations or texture changes can be detected before they become visible to the human eye.

Real-time Monitoring using Drones and IoT: Drones equipped with AI-powered cameras can scan large farms and detect diseased plants.

- **IoT sensors** collect data (e.g., temperature, humidity) and feed it into deep learning models to predict conditions favorable for disease outbreaks.

Transfer Learning for Faster Model Training: Instead of training models from scratch, pre-trained deep learning models (e.g., MobileNet, EfficientNet) can be fine-tuned for specific crop diseases, reducing training time and improving accuracy**.**

Mobile Applications for Farmers: Smartphone apps powered by deep learning allow farmers to take a picture of a leaf and get an instant diagnosis**.**

- **Example**: PlantVillage app (powered by TensorFlow) helps in real-time disease identification.

Reduction in Chemical Usage: By pinpointing affected plants, farmers can target pesticide use instead of spraying the entire field, reducing costs and environmental impact.

1.7 Conclusion

Deep learning is transforming agriculture by providing automated, accurate, and real-time disease detection. By integrating AI with smart farming technologies, farmers can increase productivity, reduce losses, and ensure sustainable agricultural practices.

2. LITERATURE REVIEW

A thorough analysis of the literature was carried out by examining research articles from reputable magazines such as IEEE and ACM. In order to chart the development of deep learning applications in agriculture, the evaluation was organized chronologically. Research has shown that deep learning models—in particular, CNN-based architectures like AlexNet, VGG, ResNet, and EfficientNet—have achieved remarkably high classification accuracy for plant diseases.

2.1 Review of Existing Work

Muhammad, Nazeer, et al. [1] state that pre-trained architectures include VGG19 CNNs and AlexNet. It was able to extract characteristics from the provided data with painstaking accuracy. Following the extraction of the ideal subset of features, the convolutional neural network feeds those features to several classifiers, such as the Artificial Neural Network, Support Vector Machine, Probabilistic Neural Network, K-Nearest Neighbour, and Fuzzy Logic. A self-gathered dataset created throughout the augmentation process was used to validate the suggested approach.

Plant leaves may be classified as either healthy or ill using a computationally efficient technique, according to Bhagat and Monu et al. [2]. If the leaves are determined to be unwell, plant leaf diseases can be identified. The Support Vector Machine was optimized using the Grid Search approach. A system for identifying and categorizing plant diseases was being developed using SVM as an algorithm. This strategy's practical approach to disease identification with minimal computing effort will benefit the farmers.

Shruthi, U., and V. Nagaveni et al. [3] explain the procedures of a general system for diagnosing plant illnesses and a comparison analysis using machine learning classification algorithms. This article compares five different machine-learning classification methods for plant disease identification. Unlike other classifiers, the SVM classifier was used by several writers to classify diseases.

Froldi, Guglielmina, et al. [4] found that aloin and aloe-emodin were only partially responsible for the antiglycation and antiradical effects of the methanolic and hydroalcoholic leaf extracts. While neither extract is cytotoxic, these two anthraquinones have somewhat negative effects on cell survival.

To find out how aloe vera and its two components—carbohydrates and polypeptides—work to reduce diabetes in rats with streptozotocin-induced diabetes, Babu, Spoorthy N. A proteomics approach was employed by et al. [5]. Over the course of three weeks, several groups of rats were given aloe vera extract, carbohydrate fraction, and peptide/polypeptide fraction. Aloe vera and its two components helped diabetic rats' insulin and glucose levels return to normal. A proteome analysis was performed on the rats' plasma following the removal of albumin and IgG.

A naturally occurring substance that is now widely utilized in the cosmetology industry is aloe vera, according to Gauri Deshpande and Pratiksha Shinde et al. [6]. Controlled studies are required to ascertain its true effectiveness, despite the fact that it has several indications for usage. This article provided an overview of the characteristics, mode of action, and therapeutic applications of the aloe vera plant. Aloe vera has been used and recognized for millennia due to its advantages for skin, health, and appearance. Two thousand years ago, Greek scientists believed aloe vera to be a cure-all. The Arabic term "Alloeh," which means "shining bitter material," is the source of the name Aloe vera, while the Latin word "vera" means "truth." Today, dermatologists utilize the aloe vera plant for many purposes; the Egyptians called it "the herb of immortality."

According to Ahmmed, Md., a new record for the Alternariaalternata-caused leaf spot disease of Aloe vera was set in Bangladesh. Sabbir and associates [7]. Using behavioral characteristics and genomic studies, the fungus species was identified. Maximum Software analyzed the ITS sequencing data and produced the phylogenetic tree. A. Probability supports alternata's inclusion in the Pleosporaceae family. The effects of various growth medium, temperature, and light on the fungi's nutritional requirements were then assessed. Large, rough patches were seen on the surface of the brownish-colored leaves of aloe vera plants afflicted with fungus. On PDA medium, the fungus had a grayish to black color with a lighter border. There were tiny, grayish aerial hyphae scattered throughout the colonies.

Based on existing pharmacological data, the majority of study focuses on antibacterial capabilities, protection of the skin and digestive system, and anti-cancer activity (Sánchez, Marta, et al., 2008). In vitro and in vivo investigations are the most recent. Only aloe vera has undergone clinical testing; isolated chemicals have not. Investigating the therapeutic effects of important metabolites in a variety of diseases and human circumstances would thus be fascinating. The encouraging outcomes of these fundamental research investigations motivate

more clinical trials to explore the therapeutic application of aloe vera and its primary components, especially for diabetes, cancer, and bone protection.

According to Uda, M. N. A., et al. [9], the fungus Pyriculariaoryzae is the cause of the severe and widespread rice blast disease that harms rice worldwide. Even regions with mild infection can lose up to 50% of their crop, and the fungus kills enough rice every year to feed an estimated 60 million people. As a result, using herbal plants to cure plant diseases was an option. The rice blast disease may be treated with herbal extracts from plants including aloe vera, citrus hystrix, zingiberofficinale, and Sabah snake grass. This study aimed to assess the antibacterial activity of these four herbal plants against the rice blast disease-causing fungus Pyriculariaoryzae.

According to Aziz, Sumair, and others [10]. Plant disease classification and identification using a computer vision framework. The proposed method extracts Local Tri-directional Patterns (LTriDP) from images of plant leaves of different classes. LTriDP features efficiently extract discriminant information and represent each class with fewer dimensions. For classification, multiclass support vector machines (SVM) were employed. The trials are conducted using the tomato leaf dataset, which has five different classifications.

According to Tiwari, Divyansh, and others [11], a model was introduced that extracts the relevant characteristics from the dataset by fine-tuning (transfer learning) pre-trained models like VGG19. The findings were then viewed using a variety of classifiers, although logistic regression performed the best in terms of classification accuracy across the test dataset.

According to Shrivastava, V. K., Pradhan et al. [12] concentrated on the identification and classification of multiple diseases in rice plants using a framework that combined CNN architecture and SVM. The system was evaluated using a dataset that included 619 images of rice plant leaves, representing all four disease classes. When the accuracy was assessed for different ratios of training and testing datasets, a maximum accuracy of 91.37 percent was reached. The early identification and categorization of crop plant diseases is one of the most crucial elements of agricultural operations. Farmers experience large financial losses each year as a result of disease infestation.

Uğuz, Sinan, and NeseUysal [13] used a data collection that contained 3400 samples of olive leaves, including healthy leaves, to detect Aculusolearius and Olive Peacock Spot Diseases, which were common diseases of olive trees in Turkey. Along with the VGG16 and VGG19

designs, they also used transfer learning approaches to our proposed CNN architecture. Both Aculusolearius and Olive Peacock Spot, which are often seen in Turkey, were classified using deep convolutional neural network models. We collected the data set for the study ourselves, which consists of 3400 samples of healthy leaves. The data was divided into training and test sets using the fivefold cross-validation technique. The classification accomplishments of three different CNN models—the proposed CNN model, Transfer learning-VGG16, and Transfer learning-VGG19—were compared using training datasets.

Ramesh, Shima, et al. [14] state that the four stages of implementation are dataset construction, feature extraction, classifier training, and classification. The datasets of the two leaf kinds are integrated and trained using Random Forest after the photos of healthy and ill leaves have been classified. Utilize the Histogram of an Oriented Gradient (HOG) to extract features from a picture. Overall, plant disease may be widely detected by employing machine learning to train massive publicly accessible data sets.

According to Varghese, BaizelKurian, et al. [15], they have created a system that uses CNN and image processing to identify plants. An in-depth knowledge of plants is necessary to identify novel or uncommon species. The differences in leaf features facilitate comparisons between different plant species.

According to Martínez-Sánchez, Ascensión, et al. [16], a significant number of fatty acids were found, with polyunsaturated unsaturated fatty acids (PUFAs) having the highest concentration of -linolenic acid and -linoleic acid at a ratio around one. Stearic acid and palmitic acid were the two most prevalent saturated fatty acids. The third stage of development has the lowest quantities of fatty acids. Aloe vera flowers contain bioactive chemicals that may find use in the nutraceutical, cosmetic, pharmacological, and culinary sectors. Depending on the compound of interest, it may be advantageous to harvest flowers while they are just beginning to mature. This minimizes the energy requirements of the flowers from the plant and promotes plant development.

According to Daniya, T., and S., the use of several machine learning and image processing approaches lessens issues in the agriculture industry. [17] Vigneshwari. This review's primary objective is to detect rice plant illnesses utilizing a range of machine learning and image processing methods in conjunction with photos of sick rice plants. Key machine learning (ML) and image processing techniques for identifying and categorizing plant diseases are

covered. Numerous classification methods, such as the K-Nearest Neighbour Classifier (KNN), Support Vector Machine (SVM), Genetic Algorithms (GA), and Probabilistic Neural Networks (PNN), are used in agricultural research. Choosing a categorization method is an important duty as the input data might affect the quality of the final product.

According to Kulkarni, Pranesh, et al. [18], computer vision and machine learning are used in a novel and effective method for detecting agricultural diseases. Manually searching for plant diseases is difficult. It requires a great deal of time and work, as well as knowledge of plant diseases. Thus, plant diseases may be identified using machine learning models and image processing. In this work, we have shown how to diagnose plant diseases using photos of the injured plants' leaves. A subset of signal processing known as "image processing" is capable of extracting relevant information or visual characteristics. They examined a range of image metrics or characteristics to determine the most accurate way to diagnose various plant leaf diseases. In the past, plant diseases were identified by visual inspection of the leaves or specialist chemical procedures. On large farms, this is particularly costly because it requires a huge team of specialists and constant plant monitoring.

According to Bhagat, Monu, et al. [19], if plant leaves are found to be ill, there is a computationally effective method for identifying the illness and classifying the leaves as either healthy or unhealthy. Support Vector Machines, the foundation of this classification method, are optimized through the use of the Grid Search methodology.

Table 2.1: Analysis of Papers (25 to 55) for Plant Leaf Diseases

S. No.	Author and Citation	Publisher (Year)	Technology Used and Findings	Limitations
1	DivyanshuTirkey, Kshitiz Kumar Singh, ShrivishalTripathi [25]	Smart Agricultural Technology, 2023	YoloV5, InceptionV3, and CNN	Computation time high More complex
2	R. Sujatha, J.M. Chatterjee, N.Z. Jhanjhi [26]	Elsevier, 2021	Comparison of AlexNet and VGG16 for leaf disease detection	High computational cost for training
3	A. Lakshmanarao, M.R. Babu [27]	IEEE, 2021	ConvNet architecture for disease classification	Limited dataset size reduces generalization

4	V.S. Dhaka, S.V. Meena, G. Rani [28]	MDPI, 2021	Survey of deep learning methods for leaf disease prediction	Difficulty in selecting best DL architecture
5	A.A. Ahmed, G.H. Reddy [29]	MDPI, 2021	Mobile-based CNN model for detection	Dependency on high-quality image inputs
6	M.M. Islam, M.A.A. Adil [30]	Elsevier, 2023	DeepCrop web application with DL models	Lacks scalability for real-time applications
7	N. Krishnamoorthy, L.V.N. Prasad [31]	Elsevier, 2021	DNN with transfer learning for rice disease prediction	Transfer learning requires large labeled data
8	P.E. Rubini, P. Kavitha [32]	IEEE, 2021	DL model for early prediction of leaf diseases	Early-stage prediction accuracy is moderate
9	S. Nandhini, K. Ashokkumar [33]	IEEE, 2021	CNN-based methods for tomato plant disease detection	Focuses on tomato diseases only
10	M. Akila, P. Deepan [34]	Academia.edu, 2018	DL for automated plant disease classification	Limited disease classes considered
11	J. Eunice, D.E. Popescu, M.K. Chowdary [35]	MDPI, 2022	NASMobileNet for crop leaf disease identification	Requires high computational resources
12	A. Pal, V. Kumar [36]	Elsevier, 2023	AgriDet framework for severity classification	Framework lacks multi-crop capability
13	A. Wongchai, D. RaoJenjeti [37]	Elsevier, 2022	Farm monitoring via DL architectures	Restricted to specific farm environments
14	R. Mahum, H. Munir, Z.U.N. Mughal [38]	Taylor & Francis, 2023	Efficient DL model for potato leaf disease	Potato-specific model with less generalization
15	I. Bhakta, S. Phadikar, K. Majumder [39]	Springer, 2023	Thermal imaging-based prediction for rice disease	Thermal imaging needs specialized equipment
16	C. Jackulin, S. Murugavalli [40]	Elsevier, 2022	Comprehensive ML and DL model comparison	Comparison is theoretical without implementations
17	G. Endait, A. Nikam, P. Bhorkade [41]	IEEE, 2023	Ensembling methods for hybrid	Limited scalability for large farms

			disease prediction	
18	M.M. Islam, M.A. Talukder, M.R.A. Sarker [42]	Elsevier, 2023	Smart web app for cotton disease detection	Focused on cotton, lacks crop diversity
19	D. Dhabliya [43]	Yashika Journals, 2021	Integrated optimization model for paddy diseases	Model not validated on real-world farms
20	G. Delnevo, R. Girau, C. Ceccarini[44]	IEEE, 2021	Social IoT and DL for sustainable agriculture	Data dependency limits IoTapplication
21	Ü. Atila, M. Uçar, K. Akyol [45]	Elsevier, 2021	EfficientNet for plant leaf classification	EfficientNet may overfit small datasets
22	Z. Liu, R.N. Bashir, S. Iqbal [46]	IEEE, 2022	IoT-based ML for tea leaf blister blight prediction	Limited to tea plants, lacks versatility
23	J. Rashid, I. Khan, G. Ali [47]	MDPI, 2021	Multi-level DL for potato disease detection	Dataset size affects prediction consistency
24	J.A. Wani, S. Sharma, M. Muzamil [48]	Springer, 2022	Hybrid computational methods for disease detection	Challenges in multi-class disease detection
25	H. Amin, A. Darwish, A.E. Hassanien [49]	IEEE, 2022	End-to-end DL model for corn disease classification	Corn-specific application limits usage
26	J. Kotwal, R. Kashyap, S. Pathan [50]	Elsevier, 2023	Traditional to DL for apple diseases	Focused on apple, lacks crop diversity
27	T.V. Reddy, K.S. Rekha [51]	IEEE, 2021	Deep Leaf Disease Prediction Framework	Transfer learning-based methods only
28	F. Arshad, M. Mateen, S. Hayat [52]	Elsevier, 2023	PLDPNet for potato disease prediction	Potato-specific framework, lacks diversity
29	M.E.H. Chowdhury, T. Rahman, A. Khandakar [53]	MDPI, 2021	Reliable leaf disease detection using DL	Heavy reliance on labeled datasets
30	S. Pawar, S. Shedge, N. Panigrahi [54]	IEEE, 2022	Multi-plant disease detection framework	Focused on small plant datasets
31	V. Pallagani, V.	IEEE, 2019	DCrop for smart	High computational

	Khandelwal, B. Chandra [55]		agricultural disease detection	needs for training large datasets

2.2 Problem Definition

- This research aims to create a deep learning-based system capable of precisely identifying and categorizing diseases that impact plants.
- Succulent in nature and often grown for its therapeutic qualities, prone to a number of diseases that can affect its quality and rate of growth.
- Conventional illness detection and classification techniques depend on human experts doing manual inspections, which can be error-prone and time-consuming.
- In order to automate and raise the precision of illness diagnosis in plants, it is necessary to take advantage of deep learning techniques.

2.3 Research Gap

- **Improving Model Generalization:** Deep learning models for plant disease prediction often struggle to generalize across diverse crops, species, and environments.
- **Early Disease Detection:** Current models struggle to detect diseases in their early stages when symptoms are subtle and hard to distinguish.
- **Data Imbalance:** Limited annotated datasets for some plant diseases cause model bias and poor performance on underrepresented diseases.
- **Real-Time Detection on Edge Devices:** Deep learning models need high computing power, making real-time disease detection on devices like smartphones and drones challenging.
- **Improved Model Explainability:** Many deep learning models for plant disease detection lack transparency, making it hard for farmers to trust and use them.

3. PRESENT STUDY

Agriculture continues to play a significant role in the production of food worldwide, and crop health is crucial to preserving food security. In recent years, one of the largest issues facing farmers worldwide has been plant diseases. Crop illnesses caused by bacteria, viruses, fungi, and other pathogens can cause significant production and quality losses as well as crop destruction. Early and accurate identification of plant diseases is necessary to reduce these losses. Plant diseases were once manually diagnosed by experts, but this process is labor-intensive and error-prone. The development of deep learning models for plant disease prediction, which provide scalable, accurate, and efficient alternatives, has totally changed this process.

Deep learning, a branch of machine learning (ML), uses artificial neural networks (ANNs) with several layers of nodes to automatically find complex patterns and representations in data. Because of their ability to handle massive amounts of data and produce predictions using learned qualities, deep learning models have become more and more popular across a range of industries, including agriculture. One of the most intriguing applications of deep learning in agriculture is the prediction and identification of plant leaf diseases from photographs. Using deep convolutional neural networks (CNNs) and other state-of-the-art deep learning architectures, this technique automates the detection of plant diseases from leaf pictures, significantly increasing the speed and accuracy of diagnosis.

3.1 The Importance of Plant Leaf Disease Detection

Plant diseases are caused by a variety of pathogens and often manifest as lesions, wilting, discoloration, and patches on plant leaves. Controlling the spread of disease and preventing significant crop losses depend on early recognition of these indications. Laboratory testing, field surveys, and expert knowledge were formerly used to identify plant diseases; however, these methods were time-consuming and often impractical in large-scale agricultural settings.

Advances in digital image processing, machine learning, and deep learning have enabled automated systems for plant disease diagnosis. The advancement of remote sensing technology, drones, and smartphone applications has made it feasible to collect large amounts of visual data. This facilitates the use of machine learning techniques to the detection of plant diseases. However, the performance of the deep learning models used to analyze the images

is what mostly determines how successful these techniques are. Deep learning algorithms are increasingly being used by academics due to their ability to automatically extract information from photographs, understand complex patterns, and provide extremely accurate predictions.

3.2 Key Deep Learning Models for Plant Leaf Disease Prediction

The topic of plant disease identification has been tackled using a variety of deep learning architectures. Convolutional Neural Networks (CNNs) are the most well-known of these, but a number of other models and approaches have also been investigated. Here are a few of the most well-known models and methods used to predict plant leaf diseases.

• CNNs: The most popular deep learning models for image classification, CNNs are especially well-suited for jobs involving highly visual data, such as plant disease diagnosis. Convolutional, pooling, and fully linked layers are among the several layers that make up CNNs. The ability of convolutional layers to automatically learn hierarchical feature representations from input images is essential for identifying patterns in plant leaves, including lesions, discolorations, and structural alterations. In numerous studies, CNNs have been successfully used to classify plant diseases from leaf images. For instance, a CNN model was trained to recognize illnesses in 14 distinct plant species in a study conducted by Mohanty et al. (2016). Using a dataset of 54,000 photos, the algorithm successfully distinguished between healthy and sick leaves with an astounding 99.35% accuracy rate. This outcome showed how effective CNNs are in identifying minute patterns in plant leaves that might be signs of illness.

CNNs are frequently used because of their capacity to process vast, high-dimensional picture data, which makes them perfect for jobs involving the identification of plant diseases. With great success, several CNN variants—including AlexNet, VGGNet, ResNet, and InceptionNet—have been optimized and modified for the categorization of plant diseases. Another research, for example, classified maize illnesses using a modified version of VGG16, with a 97.1% accuracy rate.

• Transfer Learning: This method involves fine-tuning a previously trained model using a fresh, domain-specific dataset. In plant disease identification, where labeled datasets may be scarce, transfer learning is very helpful. Researchers can get better results with relatively small datasets by using models that have previously been pre-trained on huge

datasets, like ImageNet. This is because the models already include broad feature representations that can be tailored to the job at hand. Ferentinos (2018), for instance, used a pre-trained VGG16 model using transfer learning to categorize plant illnesses in a dataset of tomato leaf photos. To identify certain illnesses like bacterial spot, late blight, and early blight, the model was refined. When compared to training from scratch, the study discovered that transfer learning greatly enhanced the model's performance, proving its usefulness in the diagnosis of plant diseases with little labeled data.

Similar to this, transfer learning has been used to the classification of plant diseases using pre-trained models like as ResNet, InceptionV3, and DenseNet, which have demonstrated notable gains in classification accuracy and generalization to unknown datasets. For deep learning models to perform better in agricultural applications, transfer learning remains a crucial approach.

• Long Short-Term Memory Networks (LSTMs) and Recurrent Neural Networks (RNNs): LSTMs and RNNs are networks that are intended for sequential data, but CNNs operate best with spatial data. When temporal or sequential data is available, RNNs—more especially, LSTMs—have been investigated for the prediction of plant diseases. This might involve tracking a disease's development over time in agricultural settings, where time-series data gathered from several photos can yield more precise forecasts of the disease's condition.

Through the analysis of leaf picture sequences across time, LSTMs have been used to forecast the course of plant illnesses, including fungal infections. Better decision-making for disease prevention and management is made possible by LSTMs' ability to identify long-term relationships in data, which can reveal insights into how illnesses arise and spread.

The family of deep learning models known as Generative Adversarial Networks (GANs) is capable of producing synthetic data. The discriminator and generator networks, which make up a GAN, are trained in tandem using a game-theoretic framework. The discriminator assesses the synthetic pictures produced by the generator and learns to differentiate between authentic and fraudulent images.

By creating artificial pictures of plant leaves with different illnesses, GANs have been

utilized to enhance limited datasets in the context of plant disease diagnosis. This aids in tackling the problem of sparsely annotated data, particularly for underrepresented plant species or uncommon illnesses. GANs enhance the resilience and generalization of deep learning models by expanding the quantity and variety of the training dataset. For instance, Sunder et al. (2020) enhanced an existing dataset for CNN training by using GANs to create artificial pictures of sick plant leaves. It was demonstrated that this method enhanced the disease identification model's performance, especially for uncommon illnesses with insufficient real-world photos.

• Hybrid Models and Ensemble Learning: To improve prediction accuracy in plant disease detection, hybrid models that include many deep learning architectures have also been investigated. Some methods, for instance, integrate CNNs with random forests, support vector machines (SVMs), or decision trees. With the help of these ensemble learning approaches, many models can focus on various elements of the job, and the final prediction is determined by the consensus of the multiple models. By decreasing the possibility of overfitting and enhancing the model's capacity to generalize to new data, ensemble approaches can increase classification accuracy. CNNs are utilized to extract visual information from pictures, while SVMs or decision trees are used for classification in hybrid models used in plant disease diagnosis. The accuracy and dependability of illness forecasts have been demonstrated to be enhanced by this combination of models.

3.3 Objective of Research Work

- To study and analysis of plant leaf disease prediction using deep learning.
- To design hybrid deep learning classification algorithm for detection of multiple plant leaf diseases.
- To explore and validate the proposed algorithm results with various existing algorithms.

The objective of this work to create a deep learning-based system that can identify and categorize diseases in plants. The goal of the study is to increase the efficiency and accuracy of illness diagnosis plants by combining the deep learning capabilities of CNNs with transfer learning using the Alex Net model and data augmentation approaches.

3.4 Scope of the Research work

- The proposed system for disease detection and classification in the suggested system is deep learning, namely a CNN with transfer learning based on the Alex Net model. Transfer learning facilitates the use of pre-trained models and boosts performance, whilst data augmentation strengthens the model's capacity for generalization.

- The study investigates methods to improve the model's performance, like investigating ensemble learning, experimenting with various data augmentation techniques, or further optimizing transfer learning.

- The scope is to leverage deep learning techniques to improve disease diagnosis in multiple leaf plants, contributing to more efficient disease management and improved crop health.

4. METHODOLOGY

4.1 Machine learning Life cycle

Computer systems may now learn on their own without explicit programming thanks to machine learning. However, how do machine learning systems operate? Thus, they may use the machine learning life cycle to describe it. Successful machine learning initiatives are built via a cycle called the machine learning life cycle. The main goal of the life cycle is to find a solution to the job.

The seven main steps in the machine learning life cycle are listed below in figure 4.1:

- Collecting Information
- Preparing data
- Wrangling Data
- Cross-validation (Data Analysis)
- Get the model trained.
- Examine the model.
- Implementation

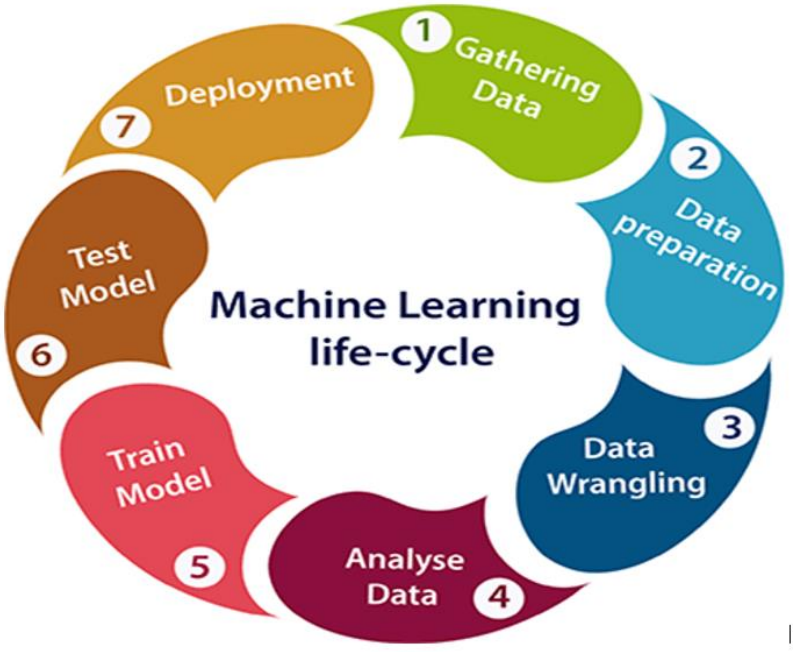


Figure 4.1: Machine Learning Life Cycle [63]

Data Gathering: Since they may gather data from a range of sources, including files, databases, the internet, and mobile devices, we must first identify the various data sources in

this stage. It is one of the most important stages of the life cycle. The effectiveness of the finished output will depend on the volume and quality of the data gathered. This stage includes the following tasks:

• Enumerate the various data sources.

• Compile data

• Integrate the data acquired from several sources.

Data preparation: We need to prepare the data for usage after it has been collected. Data preparation is the process of organizing and preparing our data for machine learning training. This stage is further subdivided by two mechanisms:

• Data exploration: It helps us identify the kinds of data we have available, along with its characteristics, formats, and other factors. A better understanding of the data leads to better results. Here you may find relationships, general trends, and outliers.

• Data preprocessing: The following step involves getting the data ready for analysis.

The process of selecting the variable to use, formatting the data suitably, and cleaning the data to get it ready for analysis in the next phase is known as "data wrangling." It involves transforming useless raw data into a format that can be utilized. Some challenges that gathered data may provide in real-world applications include the following:

• Failure to Observe Values

• Two data sets

• Inaccurate details

• Audio

Cross Validation (Data Analysis): The data is now prepared for analysis after being cleaned and prepared. A variety of techniques are used to extract characteristics. In data processing and mining, feature selection is crucial. The first attributes are condensed into a more manageable collection that yet offers enough details to differentiate across groups. This phase comprises:

The selection of analytical techniques

• Building models

Analyze the result.

Train Model: The model must be trained in order to address the problem more successfully in the next stage. We employ a variety of machine-learning techniques to train the model on datasets. They must teach a model in order to understand the numerous patterns, rules, and traits.

Test Model: We test the machine learning model once it has been trained on a specific dataset. In this step, we provide our model a test dataset to evaluate its correctness. By testing the model in compliance with the project or challenge criteria, the accuracy % of the model is ascertained.

The last stage of the machine learning life cycle, referred to as "deployment," is putting the model into use in a real system. We incorporate the model into the present system if it yields accurate, timely, and planned outcomes that meet our goals. However, we will first determine whether the project is using the data to improve performance before providing it. The final report for the project is produced during the deployment phase.

4.2 CNN Architecture

Convolutional neural networks (CNNs) are one type of artificial neural network that excels at processing data with a grid-like design, such as images and videos. Its design was inspired by how the visual cortex functions in the human brain. CNN displays many layers in figure 4.2. The following are the CNN layers:

4.2.1 Convolutional Layers

- Convolution: The core of a CNN, these layers move across the input data (such as an image) and identify features using learnable filters (kernels). In this procedure, a piece of the input is multiplied element-wise by the filter, and the outputs are then added up.
- Feature Maps: A feature map, which indicates the locations of particular patterns in the input, is the result of a convolution process. Several feature maps are produced by applying various filters to capture distinct features

4.2.2 Pooling Layers

- Downsampling: By reducing the spatial dimensions of the feature maps, pooling layers improve the network's computational efficiency and decrease its susceptibility to overfitting.
- Types of Pooling: Average pooling, which takes the average value, and max pooling, which takes the maximum value in a region, are common pooling procedures.

4.2.3 Activation Functions

- Non-linearity: By adding non-linearity to the network, activation functions help it discover intricate patterns. ReLU (Rectified Linear Unit) and its variations are popular options

4.2.4 Fully Connected Layers

- **Classification:** The CNN's fully linked layers use the feature maps that have been flattened to carry out sophisticated classification and reasoning. Conventional neural networks share similarities with these layers.

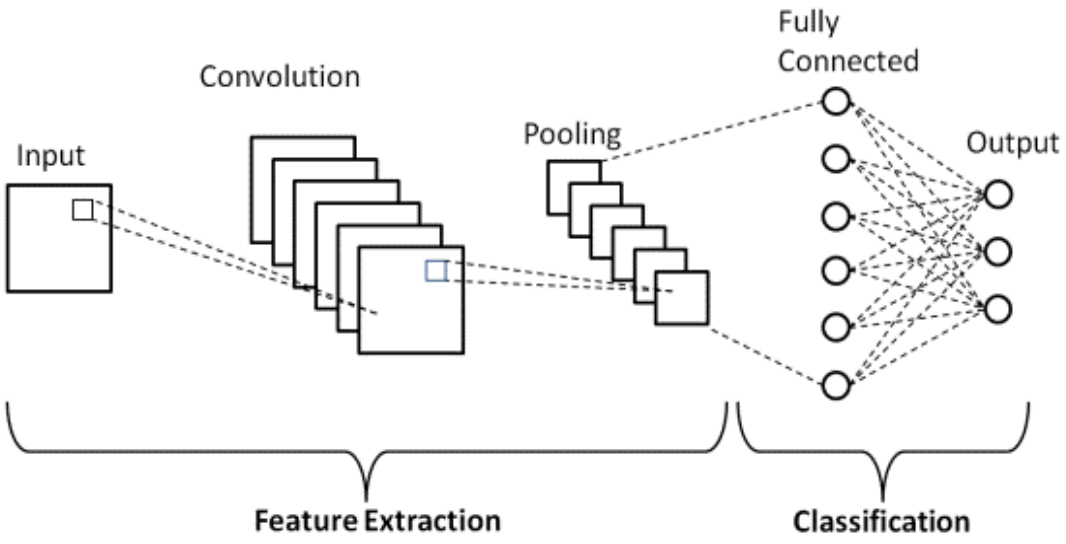


Figure 4.2: CNN Architecture [57]

4.3 System Architecture

By examining picture data, the computers are able to identify plant diseases. The system's

synthesis and real-time image data can be assessed, and any application can use them. For categorisation, the system has both a training and testing phase. The following modules figure 4.3 make up the suggested architecture of a plant disease classification and detection system.

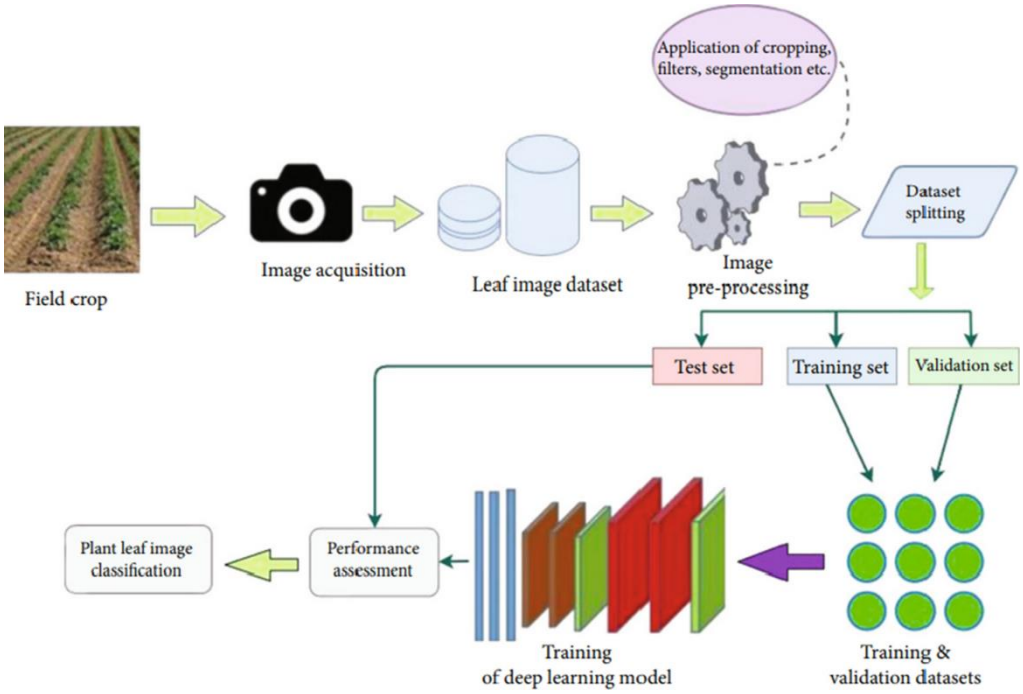


Figure 4.3: System Architecture [64]

• Data capture: picture capture is essential as the system's accuracy is largely dependent on the training picture samples. A few well-known databases in this area, such the Aloe Vera Plant Images collection, have been used by researchers.

• Preprocessing: Photos that have had distortion removed during preprocessing are easier to analyze later. Common pre-processing techniques include color space conversion, cropping, smoothing, and enhancement. The functionality of this module changes based on the quality of the image. Cropping is especially essential when taking pictures in an uncontrolled environment with complex backdrops. It can be done manually or automatically with functions.

• Segmentation: This method divides a picture into regions that are closely related to the things of interest. Healthy or contaminated samples may be easily identified using characteristics of a well-segmented picture, such as the number of histogram peaks.

• Feature extraction: To comprehend photos, color, texture, and form components are usually utilized. Color is often described using moments and histograms. Features like contrast,

homogeneity, variance, and entropy may all be linked to texture. Form, concavity, eccentricity, area, and roundness all have similar characteristics. Texture has been proven to be the most useful feature for plant disease detection systems, despite the fact that diverse datasets require a range of characteristics. A variety of techniques are used to extract characteristics.

Because the gathered photos vary in size, they must be resized to a consistent size for effective preprocessing. The RGB photos are first resized and transformed to the Hue Saturation Intensity (HSI) format. Color perception is greatly enhanced by the use of HSI color space representation. Masking is then used to eliminate the pixels. Setting a picture's pixel value to zero or another background value is known as masking. The original image is then separated into the sick region using the K-means segmentation technique. The goal of segmentation is to turn a picture's representation into a meaningful image that is simpler to examine. The best attributes from this dataset are then selected for accurate categorization via feature selection. Information Gain (IG), Principal Component Analysis (PCA), and Relief-f Attribute Evaluator (RAE) are the three feature selection techniques used in this study.

• Feature Selection: An essential stage in image processing and data mining. It finds the ideal subset of anticipated qualities from the original data. To successfully distinguish across classes, a subset of the original characteristics is chosen that retains enough information. For feature selection, a variety of search techniques may be used, such as IG, PCA, and RAE.

• Classification: Classification is an essential part of plant disease detection systems. Classification is the process of grouping images of plant leaves based on recognized diseases. Many leafplant illnesses are identified by the method using a picture. Prior to classifying or recognizing images from a test set, the classifier is trained using images from a training set. Researchers have looked into a range of deep learning methods to identify diseases in different cultures. Images showing healthy and sick leaves will be distinguished by the classifier

4.4 System Requirement and Specifications

A 500 GB hard drive, 4 GB of RAM, a GPU card, and an Intel i3 or higher processor are required for this application. Software requirements include Python 3.6 or later as the programming language, with Python serving as the main tool, and an operating system such as Windows 10 or higher.

5. DATA ANALYSIS/DISCUSSION

Data forms the cornerstone of advancements in deep learning models, especially in agricultural applications such as plant leaf disease prediction. The availability of high-quality, annotated datasets significantly influences the performance and generalization of models. In the context of plant disease detection, datasets need to encompass a wide range of disease categories, environmental conditions, and plant species. This document explores the **PlantVillage dataset**, one of the most widely used datasets in this domain, along with relevant supplementary datasets. Each dataset's characteristics, composition, and application potential are described in detail to emphasize their importance in deep learning-based plant disease prediction.

5.1 PlantVillage Dataset

The **PlantVillage dataset** [58] is a publicly available dataset specifically curated for research in plant disease identification and classification. It is one of the largest annotated image repositories for plant health studies, containing images of plant leaves affected by various diseases as well as healthy leaves. Developed by researchers at Pennsylvania State University, this dataset serves as a benchmark for many machine learning and deep learning models in agricultural applications.

Dataset Composition: The dataset comprises over 50,000 labeled images of plant leaves, covering 38 classes of plant-disease combinations. These images are captured under controlled conditions, ensuring high resolution and clarity. The dataset includes:

- **Healthy Leaves**: Images of disease-free leaves from different plant species.
- **Diseased Leaves**: Images of leaves showing symptoms of specific diseases.

Key Features:

- **Diversity of Plant Species**:
    - The dataset includes images from 14 plant species such as tomatoes, potatoes, grapes, apples, and corn.
    - These species represent staple crops globally and are vital for food security.

- **Variety of Diseases**:
    - Common diseases like late blight, bacterial spots, and powdery mildew are represented.
    - This diversity enables the development of robust models capable of handling multiple diseases.
- **Data Augmentation Potential**:
    - While the dataset provides images captured under controlled conditions, its versatility makes it suitable for data augmentation techniques. Techniques like rotation, flipping, and cropping can artificially expand the dataset's size, helping deep learning models generalize better to field conditions.

Use in Deep Learning: The PlantVillage dataset has been instrumental in training Convolutional Neural Networks (CNNs) for plant disease detection. Models such as ResNet, AlexNet, and InceptionNet have been fine-tuned using this dataset, achieving high levels of accuracy in identifying diseases. For instance, researchers using the dataset have reported classification accuracies exceeding 98% when applying transfer learning techniques. Additionally, the dataset's uniformity in image backgrounds makes it an ideal starting point for experiments requiring minimal preprocessing.

Limitations: While the PlantVillage dataset is invaluable, it has some limitations:

- **Controlled Environment**: All images are taken in laboratory settings, lacking the variability introduced by field conditions such as lighting, occlusions, and complex backgrounds.
- **Class Imbalance**: Some diseases are underrepresented, which may lead to model bias during training.

5.2 Relevant Supplementary Datasets

5.2.1 Flavia Leaf Dataset

The **Flavia Leaf Dataset** [59] is another widely used dataset for plant classification tasks. Unlike the PlantVillage dataset, which focuses on plant diseases, the Flavia dataset primarily

aims at plant species identification.

Dataset Composition:

- Contains 1,900 images representing 32 species of plants.
- Images are captured in controlled environments, focusing on the geometric and structural features of leaves.

Applications: While not disease-focused, the Flavia dataset complements the PlantVillage dataset by providing insights into plant taxonomy. It can help develop hybrid models that combine species identification and disease prediction.

Limitations:

- Small dataset size compared to PlantVillage.
- Limited to species identification, making it less suitable for disease-specific tasks.

5.2.2 Mendeley Dataset: Apple Leaf Diseases

The **Apple Leaf Disease dataset [60]** hosted on Mendeley is specifically curated for apple disease prediction. This dataset addresses some limitations of the PlantVillage dataset by including real-world variability.

Dataset Composition:

- Comprises 2,300 images of apple leaves.
- Includes images captured in both controlled and natural environments.
- Features diseases like apple scab, black rot, and cedar apple rust.

Applications:

This dataset is particularly valuable for domain-specific tasks such as apple orchard disease management. It provides diverse data for developing robust models capable of handling real-world conditions.

Limitations:

- Limited number of images and disease categories.
- Primarily focuses on a single crop, reducing its generalizability.

5.2.3 Kaggle Dataset: Crop Disease Dataset

The **Crop Disease Dataset** [61] available on Kaggle serves as another valuable resource for plant disease studies. It includes images from various crops and disease categories, offering a diverse set of training data.

Dataset Composition:

- Features over 87,000 images across 25 different crops and 58 disease classes.
- Includes real-world field images, offering variability in lighting, angles, and backgrounds.

Applications: This dataset is suitable for developing and testing deep learning models under challenging field conditions. It has been used to benchmark models like DenseNet and EfficientNet.

Limitations:

- Requires extensive preprocessing due to variability in image quality.
- May include noise in the form of mislabeled or poorly captured images.

5.3 Discussion

The synergy of datasets like PlantVillage, Flavia, Mendeley's Apple Leaf Disease dataset, and Kaggle's Crop Disease Dataset allows researchers to train models with enhanced accuracy and generalization. By integrating these datasets, researchers can:

- **Improve Robustness**:
    - Combining datasets increases the diversity of training data, making models more robust to unseen conditions.

- **Address Dataset Limitations**:
    - Supplementing the PlantVillage dataset with real-world datasets like the Apple Leaf Disease dataset mitigates its lack of field variability.
- **Enhance Transfer Learning**:
    - Using pre-trained models on larger, diverse datasets like the Kaggle Crop Disease Dataset before fine-tuning on smaller datasets improves performance.
- **Facilitate Domain Adaptation**:
    - Leveraging datasets with varied environmental conditions enables models to adapt better to specific agricultural scenarios.

5.4 Summary

High-quality datasets like PlantVillage and supplementary datasets such as Flavia, Mendeley's Apple Leaf Disease dataset, and the Kaggle Crop Disease Dataset are instrumental in advancing deep learning models for plant disease prediction. Each dataset brings unique strengths and limitations, making their combined use essential for building comprehensive and reliable models. These datasets not only support academic research but also enable practical applications in precision agriculture, helping farmers detect and mitigate plant diseases efficiently.

6. CHALLENGES AND LIMITATIONS

The suggested model details many CNN-based feature extraction and selection methods. Convolutional neural networks (CNNs) develop hierarchical representations that are resilient to translation, rotation, and scaling by utilizing convolutional layers to extract parameters from input pictures and pooling layers to minimize their size. The feasibility of the custom-developed CNN model was tested in a number of trials. Furthermore, other plant species may be studied using the same methods as this study work to create comparable models for disease prediction and detection in plants. More photos of sick and pest-ridden plants, as well as healthy plants in a variety of settings and lighting, can be added to the dataset in subsequent studies. The models' capacity to generalize to photos of plants, which are currently unknown, can be enhanced in this way.

While deep learning models have shown great promise in plant disease detection, there are several challenges and limitations that researchers continue to address:

- **Limited Annotated Data**: One of the biggest challenges in training deep learning models for plant disease detection is the scarcity of labeled data. Annotating large datasets of plant images can be time-consuming and expensive. As a result, many models rely on transfer learning or data augmentation to address this issue.
- **Variability in Environmental Conditions**: Environmental factors such as lighting, weather, and soil conditions can affect the appearance of plant leaves, making it difficult for models to generalize across different environments. Models trained in specific settings may struggle to perform well in new or diverse environments.
- **Model Interpretability**: Deep learning models, particularly CNNs, are often considered black boxes, meaning that it can be difficult to understand how they arrive at their predictions. In agricultural applications, where trust and explainability are critical, improving model interpretability is an ongoing area of research.
- **Handling Multiple Diseases**: Many plant species are susceptible to a wide range of diseases, and a single leaf can exhibit symptoms of multiple diseases simultaneously. Designing models that can handle such complex scenarios and correctly identify multiple diseases in a single image is an ongoing challenge.

7. WORK PLAN FOR RESEARCH

The proposed research work is structured into six distinct phases, spanning 24 months starting from February 2025. Each phase is carefully planned to ensure systematic progression and timely completion of the objectives. Below is a detailed explanation of each phase:

Phase 1: Literature Survey (January 2025 – May 2025)

This phase focuses on conducting an extensive review of existing research to establish a strong foundation for the proposed work. The literature survey will involve studying research papers, journals, books, and technical reports related to plant leaf disease detection, deep learning, and machine learning methodologies. Special emphasis will be placed on datasets like the PlantVillage dataset and other similar repositories to understand their structure, utility, and limitations.

Phase 2: Study of Existing Techniques (June 2025 – August 2025)

This phase involves an in-depth analysis of existing techniques for plant disease detection. Techniques such as CNNs, transfer learning, GANs, and ensemble learning will be critically evaluated for their strengths and weaknesses. The study will include hands-on experiments with pre-trained models like ResNet, AlexNet, and InceptionNet, focusing on their applicability to plant leaf disease datasets. Key outcomes of this phase will be the identification of methods to overcome limitations in current techniques and the selection of a baseline approach for the proposed system.

Phase 3: GUI and Database Design (September 2025 – December 2025)

This phase focuses on designing a user-friendly graphical user interface (GUI) and organizing the database. The GUI will be developed to enable researchers and end-users (e.g., farmers or agronomists) to upload plant leaf images for disease detection. The database design will involve structuring training and testing datasets, integrating images from PlantVillage and other sources, and preprocessing them to standardize dimensions, formats, and quality. Data augmentation techniques such as rotation, flipping, and cropping will also be applied.

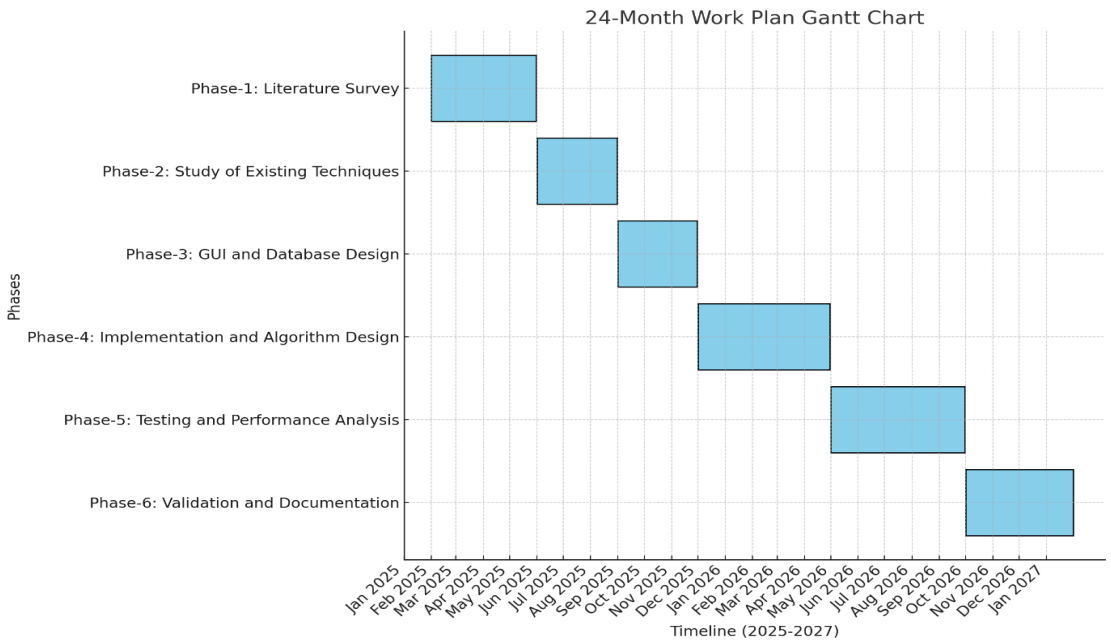


Fig 7.1: Work Plan of Proposed Work

Phase 4: Implementation and Algorithm Design (January 2026 – May 2026)

The core implementation phase will focus on developing a hybrid deep learning algorithm. This algorithm will leverage CNNs for feature extraction and incorporate transfer learning for enhanced performance on limited datasets. Techniques like GANs may be employed for synthetic data generation to augment the dataset. This phase will also involve integrating the model into the GUI for real-time predictions.

Phase 5: Testing and Performance Analysis (June 2026 – October 2026)

This phase involves rigorous testing of the proposed system using diverse datasets, including PlantVillage and real-world samples. Performance metrics such as accuracy, sensitivity, specificity, and F1-score will be used to evaluate the system. Comparative analysis will be conducted against existing methods to validate the effectiveness of the proposed approach.

Phase 6: Validation and Documentation (November 2026 – January 2027)
The final phase focuses on validating the system's usability, scalability, and robustness. Feedback from domain experts and potential users will be collected. This phase also includes compiling all research findings, creating detailed documentation, and preparing the final thesis for submission.

8. REFERENCES

[1] Muhammad, Nazeer, et al. "Severity recognition of multiple leafdiseases using AI in tensor flow domain." CMC-Computers Materials & Continua 66.2 (2021): 2199-2216.

[2] Bhagat, Monu, et al. "Plant leaf disease classification using grid search based SVM." 2nd International Conference on Data, Engineering and Applications (IDEA). IEEE, 2020.

[3] Shruthi, U., V. Nagaveni, and B. K. Raghavendra. "A review on machine learning classification techniques for plant disease detection." 2019 5th International conference on advanced computing & communication systems (ICACCS). IEEE, 2019.

[4] Froldi, Guglielmina, et al. "Antiglycation activity and HT-29 cellular uptake of Aloe-Emodin, Aloin, and Aloe arborescens leaf extracts." Molecules 24.11 (2019): 2128.

[5] Babu, Spoorthy N., S. Govindarajan, and Ayesha Noor. "Multiple leafand its two bioactive constituents in alleviation of diabetes–proteomic & mechanistic insights." Journal of Ethnopharmacology 280 (2021): 114445.

[6] Gauri Deshpande, Pratiksha Shinde, Saee Patil, Shweta Pawar, Prajakta Dhamdhere."Multiple leafPlant Diseases Detection" International Journal of Research in Engineering, Science and Management Volume 5, Issue 6, June 2022 https://www.ijresm.com | ISSN (Online): 2581-5792

[7] Sánchez, Marta, et al. "Pharmacological update properties of Aloe vera and its major active constituents." Molecules 25.6 (2020): 1324.

[8] Uda, M. N. A., et al. "Antimicrobial activity of plant extracts from aloe vera, citrus hystrix, sabah snake grass and zingiber officinale against pyricularia oryzae that causes rice blast disease in paddy plants." IOP Conference Series: Materials Science and Engineering. Vol. 318. No. 1. IOP Publishing, 2018.

[9] Aziz, Sumair, et al. "Image pattem classification for plant disease identification using local tri-directional features." 2019 IEEE 10th Annual Information Technology,

Electronics and Mobile Communication Conference (IEMCON). IEEE, 2019

[10] R. Meena Prakash, G. P. Saraswathy, G. Ramalakshmi, K. H. Mangaleswari, and T. Kaviya, "Detection of leaf diseases and classification using digital image processing," Proc. 2017 Int. Conf. Innov. Information, Embed. Commun. Syst. ICIIECS 2017, vol. 2018-Janua, pp. 1-4, 2018, doi: 10.1109/ICIICCS2017.8275915.

[11] A. Van Arendonk, "The development of the share of agriculture in GDP and employment May 2015," no. May, 2015.

[12] N. Kahanna and P. Solanki, "Agricultural & Horticultural Sciences," Agric. Hortic. Sci., vol. 2, no. 4, pp. 1-16, 2014.

[13] S. KC and W. Lutz, "The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100," Glob. Environ. Chang., vol. 42, pp. 181-192, 2017, doi: 10.1016/j.gloenvcha.2014.06.004.

[14] T. Challenge, W. Poverty, W. Anti, P. Programme, G. Myrdal, and M. Murphy, "World poverty,"no. Pelican, 1971.

[15] UN FAO; NitiAayog; MoAFW, "Indian Agriculture Towards 2030. Transforming Indian Agriculture," pp. 1-27, 2020,

[16] G. K. Jha, A. Suresh, B. Punera, and P. Supriya, "Growth of horticulture sector in India: Trends and prospects," Indian J. Agric. Sci., vol. 89, no. 2, pp. 314-321, 2019.

[17] P. Schreinemachers, E. B. Simmons, and M. C. S. Wopereis, "Tapping the economic and nutritional power of vegetables," Glob. Food Sec., vol. 16, no. June 2017, pp. 36-45, 2018,

[18] J. S. Dias, "World importance, marketing and trading of vegetables," ActaHortic., vol. 921, pp. 153-170, 2011.

[19] B. Chandra, K. Viswavidyalaya, and W. Bengal, "Production Performance of Chilli in India-A Statistical Approach Production Performance of Chilli in India- A Statistical Approach," no. March, 2017.

[20] D. Godoy et al., The future of food and agriculture: trends and challenges, vol. 4, no. 4. 2014.

[21] J. Tsopoe and N. Murry, "Economics of Chilli Cultivation in Wokha District of Nagaland, India," Curr. Agric. Res. J., vol. 8, no. 1, pp. 46-51, 2020, doi: 10.12944/carj.8.1.08.

[22] C. Penella and A. Calatayud, "Pepper Crop under Climate Change: Grafting as an Environmental Friendly Strategy," Clim. Resilient Agric. - Strateg. Perspect., no. March, 2018, doi: 10.5772 intechopen.72361.

[23] S. Lanka et al., "Chilli Outlook - February 2020," no. February, pp. 12-14, 2020.

[24] https://hscactus.org/resources/digest/plant-care/aloe-pests-and-diseases/#:~:text=meibomiae%2C%20this%20fungal%20infection%20causes,on%20haworthia%20and%20gasteria%20leaves.

[25] DivyanshuTirkey, Kshitiz Kumar Singh, ShrivishalTripathi,Performance analysis of AI-based solutions for crop disease identification, detection, and classification, Smart Agricultural Technology, Volume 5,2023,100238,ISSN 2772-3755,https://doi.org/10.1016/j.atech.2023.100238.

[26] Sujatha, R., Chatterjee, J. M., &Jhanjhi, N. Z. "Performance of Deep Learning vs Machine Learning in Plant Leaf Disease Detection." *Microprocessors and Microsystems* 81 (2021): 103727.

[27] Lakshmanarao, A., Babu, M. R., & others. "Plant Disease Prediction and Classification Using Deep Learning ConvNets." In *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)*, IEEE, 2021.

[28] Dhaka, V. S., Meena, S. V., Rani, G., Sinwar, D., Ijaz, M. F., & others. "A Survey of Deep Convolutional Neural Networks Applied for Prediction of Plant Leaf Diseases." *Sensors* 21, no. 14 (2021): 4749.

[29] Ahmed, A. A., & Reddy, G. H. "A Mobile-Based System for Detecting Plant Leaf Diseases Using Deep Learning." *AgriEngineering* 3, no. 3 (2021): 478–493.

[30] Islam, M. M., Adil, M. A. A., Talukder, M. A., & others. "DeepCrop: Deep Learning-

Based Crop Disease Prediction with Web Application." *Journal of Agriculture and Food Research* 2 (2023): 100091.

[31] Krishnamoorthy, N., Prasad, L. V. N., Kumar, C. S. P., & others. "Rice Leaf Diseases Prediction Using Deep Neural Networks with Transfer Learning." *Environmental Technology & Innovation* 24 (2021): 101746.

[32] Rubini, P. E., &Kavitha, P. "Deep Learning Model for Early Prediction of Plant Disease." In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, IEEE, 2021.

[33] Nandhini, S., &Ashokkumar, K. "Analysis on Prediction of Plant Leaf Diseases Using Deep Learning." In *2021 International Conference on Computer Communication and Informatics (ICCCI)*, IEEE, 2021.

[34] Akila, M., &Deepan, P. "Detection and Classification of Plant Leaf Diseases by Using Deep Learning Algorithm." *International Journal of Engineering Research & Technology* 7, no. 3 (2018): 1–6.

[35] Eunice, J., Popescu, D. E., Chowdary, M. K., &Hemanth, J. "Deep Learning-Based Leaf Disease Detection in Crops Using Images for Agricultural Applications." *Agronomy* 12, no. 1 (2022): 177.

[36] Pal, A., & Kumar, V. "AgriDet: Plant Leaf Disease Severity Classification Using Agriculture Detection Framework." *Engineering Applications of Artificial Intelligence* 114 (2023): 105027.

[37] Wongchai, A., Jenjeti, D. R., Priyadarsini, A. I., Deb, N., & others. "Farm Monitoring and Disease Prediction by Classification Based on Deep Learning Architectures in Sustainable Agriculture." Ecological Informatics 66 (2022): 101460.

[38] Mahum, R., Munir, H., Mughal, Z. U. N., Awais, M., & others. "A Novel Framework for Potato Leaf Disease Detection Using an Efficient Deep Learning Model." Human and Ecological Risk Assessment: An International Journal 29, no. 1 (2023): 123–143.

[39] Bhakta, I., Phadikar, S., Majumder, K., Mukherjee, H., & others. "A Novel Plant Disease Prediction Model Based on Thermal Images Using Modified Deep

Convolutional Neural Network." Precision Agriculture 26, no. 1 (2023): 45–67.

[40] Jackulin, C., &Murugavalli, S. "A Comprehensive Review on Detection of Plant Disease Using Machine Learning and Deep Learning Approaches." Measurement: Sensors 18 (2022): 100347.

[41] Endait, G., Nikam, A., Bhorkade, P., & others. "Hybrid Model to Predict Leaf Disease Prediction Using Ensembling Machine Learning Approach." In 2023 International Conference on Communication & Information Processing (ICCIP), IEEE, 2023.

[42] Islam, M. M., Talukder, M. A., Sarker, M. R. A., Uddin, M. A., & others. "A Deep Learning Model for Cotton Disease Prediction Using Fine-Tuning with Smart Web Application in Agriculture." Intelligent Systems with Applications 14 (2023): 200058.

[43] Dhabliya, D. "An Integrated Optimization Model for Plant Diseases Prediction with Machine Learning Model." In Machine Learning Applications in Engineering and Applied Sciences, Yashika Journals, 2021.

[44] Delnevo, G., Girau, R., Ceccarini, C., & others. "A Deep Learning and Social IoT Approach for Plants Disease Prediction Toward a Sustainable Agriculture." IEEE Internet of Things Journal 8, no. 6 (2021): 5267–5275.

[45] Atila, Ü.,Uçar, M., Akyol, K., &Uçar, E. "Plant Leaf Disease Classification Using EfficientNet Deep Learning Model." Ecological Informatics 61 (2021): 101182.

[46] Liu, Z., Bashir, R. N., Iqbal, S., Shahid, M. M. A., & others. "Internet of Things (IoT) and Machine Learning Model of Plant Disease Prediction–Blister Blight for Tea Plant." IEEE Access 10 (2022): 12345–12356.

[47] Rashid, J., Khan, I., Ali, G., Almotiri, S. H., AlGhamdi, M. A., & others. "Multi-Level Deep Learning Model for Potato Leaf Disease Recognition." Electronics 10, no. 12 (2021): 1379.

[48] Wani, J. A., Sharma, S., Muzamil, M., Ahmed, S., & others. "Machine Learning and Deep Learning Based Computational Techniques in Automatic Agricultural Diseases Detection: Methodologies, Applications,

[49] Amin, H., Darwish, A., Hassanien, A. E., &Soliman, M. "End-to-End Deep Learning Model for Corn Leaf Disease Classification." IEEE Access 10 (2022): 12657–12666. (portal.issn.org)

[50] Kotwal, J., Kashyap, R., &Pathan, S. "Agricultural Plant Diseases Identification: From Traditional Approach to Deep Learning." Materials Today: Proceedings 68 (2023): 765–774.

[51] Reddy, T. V., &Rekha, K. S. "Deep Leaf Disease Prediction Framework (DLDpf) with Transfer Learning for Automatic Leaf Disease Detection." In 2021 5th International Conference on Electronics, Communication, and Aerospace Technology (ICECA), IEEE, 2021.

[52] Arshad, F., Mateen, M., Hayat, S., Wardah, M., & others. "PLDPNet: End-to-End Hybrid Deep Learning Framework for Potato Leaf Disease Prediction." Alexandria Engineering Journal 62, no. 4 (2023): 3457–3471.

[53] Chowdhury, M. E. H., Rahman, T., Khandakar, A., Ayari, M. A., & others. "Automatic and Reliable Leaf Disease Detection Using Deep Learning Techniques." AgriEngineering 3, no. 3 (2021): 589–604.

[54] Pawar, S., Shedge, S., Panigrahi, N., & others. "Leaf Disease Detection of Multiple Plants Using Deep Learning." In 2022 International Conference on Machine Learning and Data Engineering (MLDE), IEEE, 2022.

[55] Pallagani, V., Khandelwal, V., Chandra, B., & others. "DCrop: A Deep-Learning-Based Framework for Accurate Prediction of Diseases of Crops in Smart Agriculture." In 2019 International Symposium on Smart Agriculture and Food Technologies (SMAFT), IEEE, 2019.

[56] Kaviani, Pouria, and SunitaDhotre. "Short survey on naive bayes algorithm." International Journal of Advance Engineering and Research Development 4, no. 11 (2017): 607-611.

[57] https://www.analyticsvidhya.com/blog/2022/03/basic-introduction-to-convolutional-neural-network-in-deep-learning/

[58] https://paperswithcode.com/dataset/plantvillage

[59] https://flavia.sourceforge.net/

[60] https://data.mendeley.com/datasets/9m2dcb5mmr/3

[61] https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset

[62] Thakur, S.K., Jain, B.C., Singh, S., Kurrey, D.and Chandrakala. (2023). A Comprehensive Economic Analysis of AloeVera Crop Cultivation in Bilaspur District, Chhattisgarh, India. AgroEconomist - An International Journal, 10(03): 213-218.

[63] https://medium.com/@sabihaali1/machine-learning-lifecycle-80c4a265e589.

[64] Tiwari V, Joshi RC, Dutta MK. Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images. Ecol Inform. 2021;63: 101289. https://doi.org/10.1016/j.ecoinf.2021.101289.

[65] Mohamed, Ahmed ElSawy, and Mohamed Afify. "Deep learning in plant diseases detection for agricultural crops: a survey." International Journal of Service Science, Management, Engineering, and Technology (IJSSMET) 11.2 (2020): 41-58.

[66] Meet p. shah, sougata singha, suyash p. awate, leaf classification using marginalized shape context and shape+texture dual-path deep convolutional neural network,978-5090-2175-8/17/$31.00

[67] Redha Ali, Russell Hardie, Almabrok Essa, A Leaf Recognition Approach to Plant Classification Using Machine Learning, 978-1-5386-6557-2/18/$31.00 ©2018 IEEE

[68] Shadi Alamoudi, Xia Hong, Hong Wei, Plant Leaf Recognition Using Texture Features and Semi-Supervised Spherical K-means Clustering, 978-1-7281-6926-2/20/$31.00 ©2020 IEEE Authorized licensed use limited to: University of Newcastle. Downloaded on October 18,2020 at 00:24:37 UTC from IEEE Xplore.

[69] Saleem, Muhammad Hammad, Johan Potgieter, and Khalid Mahmood Arif. "Plant disease detection and classification by deep learning." Plants 8.11 (2019): 468.

[70] Geetharamani, G., and Arun Pandian. "Identification of plant leaf diseases using a

nine-layer deep convolutional neural network." Computers & Electrical Engineering 76 (2019): 323-338.

[71] Zhu, Xiaolong, Meng Zhu, and Honge Ren. "Method of plant leaf recognition based on improved deep convolutional neural network." *Cognitive Systems Research* 52 (2018): 223-233.